

Hesitation: Thoughts, Pauses, Gestures and Speech

NATALYA V. SUKHOVA

Lomonosov Moscow State University (Russia)

Abstract

The paper deals with the phenomenon of pauses hesitation and their different aspects. People hesitate in speech due to several reasons which are not always obvious, so there is an insight into the speech production stage which is needed. Sometimes while speaking we hesitate and meanwhile produce different kinetic forms of behaviour (gestures, postures, facial expressions, body movements, head movements). At the same time those stretches of speech are charged with prosodic information of specific kind. Thus, those three aspects of hesitation and their coexistence are in the focus: cognitive, gestural and prosodic.

1/ about the multidisciplinary

- Why all those phenomena = what is the aim?
- What are the theories which could help to study all those things – many sciences
- Where to find the crossing lines? = zones

2/ material

- What
- Texts = marking (the parameters) =

3/ discourse (hesitation + discourse markers)

- (*outer meaning* = social surroundings + philosophy; *inner meaning* = linguistic analysis of constituents -- semantics of main parts (discourse markers) -- speech production (hesitation in ideas))
- The theories (see the book) – the aims

4/ applying the theories

- Laclau and Mouffe – environment of discourse = what is useful?
 - o The analysis
- Rhetoric Structure theory – what is so useful about it?
 - o The analysis

5/ purely linguistic approach (pauses)

- THINK ABOUT DIFFERENT APPROACHES
- Pauses and pauses hesitations (acoustic and perceptual data)
 - o The analysis

6/ nonverbal part (gestures)

- Kreidlin and practical approaches (cf.: McNeill)
- Gesture-speech interaction
 - o The analysis

7/ overlapping zone of investigation of these phenomena

- Speech-gesture production process
- Pierce

8/ the use of the multidisciplinary approach

- What's new we've learnt, why is it useful?

1. MULTIDISCIPLINARY AREA OF RESEARCH

1.1. The aim of the article

Nowadays the idea to research phenomena in a multidisciplinary way has already become very widespread. There is no point in disregarding it. However, it should be born in mind that on the one hand multidisciplinary approach may be fruitful, revealing new aspects of an object; on the other hand it may overlook some deep thus essential peculiarities of a phenomenon.

The aim of this work is to analyse discourse where hesitation pauses are filled with discourse markers and complexes of kinetic forms. Also the work aims at studying the usage of two methods of researching the material and at comparing the results. Two methods are a) to study the material with the help of some relevant specific to the phenomenon theories; b) to study it with a multidisciplinary approach.

Thus, the tasks to tackle are:

- 1) to look at the number of phenomena (discourse markers, pauses hesitations and kinetic forms) through the lenses of different relevant theories separately;
- 2) to compare the results of these analyses;
- 3) to investigate the complex of these phenomena through multidisciplinary theoretical tools;
- 4) to compare the results of a separate analysis with a multidisciplinary one.

This paper may seem to result in an overview of some theories and their general application to the discourse, or to give a very clear-cut picture of a set of steps a researcher should make to study various phenomena. However it is, the main incentive was to analyse a part of a discourse and to get as much information about it as possible through all available tools and to see which way would be the most productive.

1.1. The phenomena under discussion

Usually when we talk about a multidisciplinary approach, we mean that the phenomena under discussion are from different levels or even different branches of science, or simply it is not possible to describe them in the same terms, or we would like to look at them from the point of view of related sciences. In other words, what is different about them is: 1) the objects themselves; thus, 2) the metalanguage to describe them; and 3) the tools/methods/types of analysis to study them.

In the context of our research the phenomena are *discourse markers*, *hesitation pauses* and *kinetic forms*. Some notions have to be defined at that point. Discourse here is understood in a narrow sense; and it is considered to be a purely linguistic interpretation in comparison with sociological (cf.: Michel Foucault) and philosophical. Discourse markers are considered here to be «signals that the piece of text being processed is to be linked to some other piece of the text in a particular way» (P. 568. Maite Taboada. Discourse markers as signals (or not) of rhetoric relations // journal of pragmatics, 38, 2006. – pp. 567-592)

Hesitation pauses

Kinetic forms

The question arises how they all come together. The focus here is laid on the discourse where the hesitation pauses can be found. Those pauses may be filled, and usually are filled, with the discourse markers (e.g. you know, er-, well, etc.). In turn, discourse markers are concomitant with some kinetic forms, such as hand gestures, gestures of head, postures, etc. All of those aspects of speech and behaviour belong to different communicative levels (verbal and nonverbal), but nevertheless they function and influence the communicative situation jointly. This is the main argument to undertake the further research.

2. MATERIAL

2.1. Content and central parameters

The material consists of 5 monologues in English with total duration of 72.62 seconds. There is one man talking,^[1] he is Jamie Oliver, 28 years old at that time.

The monologues are marked throughout. There are some parameters which are specifically noted, they are:

- 1) discourse markers;
- 2) syntactic and hesitation pauses (the place and duration);
- 3) kinetic forms on the stretch of the pauses.

2.2. Example

For example, let us see the protocols of Episode **Jamie_5**.

1. ...too much fat | 0.71
2. oh (0.189) absolutely (0.636) | 0.91
3. you know (0.472)| 0.95
4. covered in additives and all that business | 0.45
5. a-er-[nd] (0.68) ι 0.132 you know (0.293) | 0.147
6. it should've never gone like that ι 0.045 that's totally irresponsible ||
[Surely, when those kids're walking to that school]
7. you would think | 0.46
8. paying taxes ι 0.05 that your kid | 0.509
9. would be eating something ι 0.106 that'd've been cared for ι 0.07 and thought about
| 0.447
10. er (0.412) ι 0.134 with a little bit of love and care | 0.374
11. from someone that actually knows what they're talking about | 0.326
12. and that's not happening ||

Where, *l* – shows the number of a line of a phrase; | - long syntactic pause; 0.71 – duration of a syntactic pause; (0.189) – duration of a speech unit; ι - short syntactic pause; || - final syntactic pause; [*a phrase*] – the speaker is not visible at the moment.

[1] The material is taken from the TV programme on the British Channel 4 – *Jamie's School Dinners*, first aired: UK, Channel 4, February 2005. More about it on <http://www.jamieoliver.com/tv/school-dinners>.

TABLE 1.- PROTOCOL OF KINETIC FORMS (EPISODE JAMIE_5)

Number of line	Hands	Head	Facial expressions	Body movements	Postures
2	+ [2]	+	-	-	-
3	+	-	-	-	+
5	+	+	+	-	-
10	+	+	+	-	-

3. DISCOURSE MARKERS

3.1. What is discourse?

Nowadays discourse may be considered as an essential element of the Humanities. It inevitably results in a rather vague definition of a term ‘discourse’, on the one hand, and a rather broad use of it, on the other. This situation in its turn gives the right to apply the notion to practically anything what is uttered, written or thought. Discourse encompasses the use of spoken, written and signed language and multimodal/multimedia forms of communication, and is not restricted to ‘non-fictional’ (e.g. stylistics) nor verbal (e.g. gesture and visual) materials. Although early linguistic approaches judged the unit of discourse to be larger than the sentence, phenomena of interest can range from silence, to a single utterance (such as «ok»), to a novel, a set of newspaper articles or a conversation. In that sense there are theories which interpret, explain and study that broadly discussed discourse. I am going to use some of those theories to show what they can highlight in a spoken discourse. Approaches that are commonly included under the term ‘discourse studies’ (or have overlapping concerns) include critical discourse analysis, critical linguistics, text linguistics, conversation analysis, ethnomethodology, discursive psychology, stylistics, genre studies, mediated discourse analysis, discourse theory, sociolinguistics, rhetorical analysis, argumentation theory, polyphony theory, etc.

In the focus of my attention there is a discourse (a speech stretch) with its internal and external formats, that is, in other words, its linguistic form and pragmatic meaning.

3.2. Pragmatic and sociocultural motivation of linguistic form in the discourse

There is some implicit «knowledge, contexts and principles which regulate the communication and thus influence the verbalization of a discourse» (Bergelson 2005: 9)^[3], because the formal organization of a discourse is evidently motivated by external factors. The linguistic form in this context, as M. Bergelson puts it (Bergelson 2005: 9-10) covers both *microstructure of a discourse*, namely, applied illocutional categories, structure of replies, self-corrections of a speaker, etc.; and *the macrostructure form of a discourse*, that is, peculiarities of a narrative scheme, change of narration types, overlapping of discourse modes, different discourse genres.

[2] This is a prominent kinetic form (see about kinetic prominence below).

[3] The translation is mine (NS).

Принцип *Приоритета* в отношении к значимым компонентам смысла регулирует способы оформления прагматически важной информации с помощью приоритетных стратегий. Прагматически важная информация при вербализации получает коммуникативно выделенное (более развернутое, грамматически престижное) оформление. А прагматически менее важная информация – редуцированное и ослабленное. Этот принцип основан как на стабильных (существующих в картине мира), так и на текущих, характеризующих данное дискурсивное взаимодействие, параметрах важности информации. Соответственно, будут задействованы абсолютные или же ситуативные приоритетные стратегии. Так, в культурной модели ПОХОД В ТЕАТР пьеса, игра актеров и качество постановки более важны, чем, скажем, униформа гардеробщика или цвет кресел в зрительном зале. Однако в конкретной беседе в рамках этой модели именно эти детали могут оказаться прагматически важными и мотивировать выделенное, маркированное, оформление этой информации лексическими, грамматическими, интонационными или другими пара- и невербальными средствами.

Принцип *Маршрутизации* информации основан на универсальных законах человеческого восприятия, категоризации и интерпретации информации. Для того, чтобы новая информация могла быть воспринята и понята, она должна выступать на фоне уже известного, категоризованного. В дискурсивных взаимодействиях этот принцип определяет то, как участники представляют информацию с точки зрения противопоставлений Известное ~ Неизвестное, Данное ~ Новое. Стратегии, реализующие эти противопоставления, это не только механизмы актуального членения, но и референция и номинация, артикли, интонационные контуры высказывания и многое другое.

Принцип *Линеаризации* информации связан с необходимостью разрешения объективного противоречия между принципиальной нелинейностью смысла и линейным разворачиванием информации в письменном тексте или устном дискурсе. Успехи нейрофизиологии в изучении нейронных сетей позволяют предположить, что смыслы, значения, знания организованы в не менее сложные, чем нейронные, также многомерные, сетевые структуры. В дискурсе они предстают в виде линейной

12

последовательности звуков или букв². Принцип линеаризации приводит в действие стратегии, также обычно относимые к актуальному членению, хотя в основе их лежит другая мотивация.

Действие принципов Маршрутизации и Линеаризации особенно хорошо видны в устном дискурсе. Различного вида коррекции, выражающиеся как паузы, обрывы, хезитации, возвраты, изменения первоначального плана изложения, повторы в описаниях, уточнения, есть проявления этих принципов.

Типы изложения могут быть следующим образом соотнесены с коммуникативными функциями:
 Коммуникативная функция: Информирование адресата
 Дескриптивный (описательный) тип изложения

Нарративный (повествовательный) тип изложения

Экспозиторный (объяснительный) тип изложения

Коммуникативная функция: Воздействие на адресата

Директивный (инструктирующий) тип изложения

Аргументативный (убеждающий) тип изложения

Типы изложения представляют собой формы дискурса, гораздо более однородные, чем жанры. Каждый тип изложения имеет характерные, точнее – прототипические, формальные (лексические и морфосинтаксические) признаки. Для нарратива это прошедшее время перфективных глаголов, для директива – формы императива, для аргументатива – модальные предикаты, для дескриптива – стативные предикаты, для экспозиторного изложения – наличие слов с обобщающим значением, не конкретно-референтный статус описываемых объектов.

PART II

В отличие от прагматической мотивированности дискурса, прагматические основания грамматических категорий могут показаться менее очевидными.

P.28

Поэтому, во многих случаях объяснить функциональные различия грамматических конструкций без обращения к прагматике и к дискурсу просто невозможно. Главными мотивирующим факторами этих различий являются знания (представления об «идеальной» картине мира) и контексты. Первые предопределяют существование в языке исходных, прототипических, конструкций, чья внутренняя структура мотивирована этой картиной, вторые – не-прототипических, воспроизводимых, а также – производных конструкций, чья прагматическая функция состоит в возможности отойти от прототипических представлений, возможности, мотивированной контекстом конкретного дискурсивного события.

Так, многочисленные исследования залогов и диатез в языках показали

определение границ дискурсивного сегмента с помощью границ жестов (см. Николаева (52))

с. 4

Дэвид МакНил, первым обративший внимание на такой интересный феномен, называет повторяющиеся жесты *catchments* (McNeill et al. 2000).³ Д. МакНил описывает это явление так: << Кэтчмент присутствует, когда какие-то характеристики сохраняются в двух или более (не обязательно последовательных) жестах. Смысл этого понятия в том, что повторение визуального представления в мышлении говорящего будет порождать повторяющиеся характеристики жеста. Выявляя кэтчменты, созданные определенным

говорящим, мы можем увидеть, что именно этот говорящий объединяет в более крупные фрагменты — какие понятия рассматриваются им как сходные или относящиеся к одной группе, а какие значения помещены в разные кэтчменты или изолированы, и таким образом воспринимаются говорящим как имеющие разные или менее связанные значения.>>

Однако поскольку главным вопросом в нашей работе было построение структуры всего нарратива, то этого фактора, взятого отдельно, оказалось недостаточно. Повторяющиеся жесты — не постоянное явление в процессе речи, сопровождаемой жестикуляцией. Они помогают выделить какие-то отдельные сегменты, связать их друг с другом, как в примере 1, но для построения целостной структуры их решительно недостаточно.

С.5

В нашей статье поднята пока малоизученная тема о связи структуры устного нарратива и иллюстративных жестов. На примере небольшого корпуса устных рассказов показано, как отдельные признаки жестов и положения рук могут добавлять дополнительную информацию касательно организации дискурса, состояния говорящего и процесса коммуникации. Так, изменение положения покоя рук с большой вероятностью указывает на границу сегмента — <<абзаца>> устной речи. Повторяющиеся жесты указывают на сохранение некоторой локальной темы или на возвращение референта, упомянутого ранее. Использование жестового пространства для повторной референции, по нашим наблюдениям, имеет свои особенности: в таких случаях жест напоминает не столько об объекте, сколько о целой ситуации из мира дискурса или о моменте в коммуникации.

5. HESITATION PAUSES

Pause – production – perception – the effect on a listener

Disfluencies – filled with UHM etc can signal that the listener should be ready for sth new
= here discourse-new and discourse-given info comes into the picture

Hesitation in speech can. . . um. . . help a listener understand

Martin Corley

Robert J Hartsuiker

Despite the many disfluencies that occur in spontaneous speech, most studies of the comprehension of spoken language have used idealised, fluent utterances. This owes much to the commonly held view that disfluencies are noise and present obstacles to comprehension

(Brennan & Schober, 2001, p. 275). However, some researchers have argued that disfluencies do not constitute noise at all, but are actually informative to listeners: they may provide information about the state of speakers' production systems. Specifically, certain disfluencies signal to listeners that speakers are experiencing production difficulty. Difficulty can occur at any stage of the process during planning, lexical retrieval, or the articulation of a speech plan and it has been argued that different types of disfluency signal different kinds of problems (e.g., Borteld et al., 2001).

To date, much of the evidence supporting this account of conversational disfluencies has come from corpus studies of filled pauses such as *uh*, *um*, *the* as /Di:/, and *oh* (e.g., Clark & Fox Tree, 2002; Fox Tree & Clark, 1997; Fox Tree & Schrock, 1999), and from experimental evidence gathered from speakers. For example, when asked to answer general knowledge questions, speakers tend to produce more *uhs* and *ums* before answers they are unsure of (Brennan & Williams, 1995; Smith & Clark, 1993). Moreover, *uh* appears to signal a shorter upcoming pause (and by inference, a less severe retrieval problem) than *um* (Smith & Clark, 1993), a finding borne out by corpus analyses (Clark & Fox Tree, 2002).

A number of studies examine in more detail the circumstances that might lead to a problem with retrieval.

For example, unpredictable lexical items are preceded by hesitations more often than those that are predictable (Beattie & Butterworth, 1979). There is also a well-established

correlation between disfluency and lexical frequency. For example, Maclay and Osgood (1959) examined a sample of spontaneous speech and found that pauses filled with *er* and the like were more likely to occur before open-class than (high frequency) closed-class words; Levelt (1983) showed that the frequency of colour names correlated negatively with the probability that these would be preceded by filled pauses. However, findings concerning production do not necessarily imply that disfluencies are somehow *designed* to inform listeners about the states of speakers' production systems: they could simply be a by-product of the speech production process.

Although a number of researchers have suggested that disfluency may play a focusing role (e.g., Brennan & Schober, 2001; Fox Tree, 1995, 2001; Fox Tree & Clark, 1997; Fox Tree & Schrock, 1999), there is a possibility that the faster reaction times after *um* in these experiments are an artefactual result of entropy (assuming that participants in this experiment were able to gain an implicit understanding of the experimental design³) : consider having heard the words *now press the button for the* . . . In cases where there is no *um*, there are still 3 possibilities at this point: *um*, or one of two target words. On the other hand, an *um* signals that the next constituent *must* be a target word, and participants may be able to use this information strategically to prepare a response. A third experiment (currently underway) addresses this problem by contrasting *now press the button for the um* . . . with *now press the um button for the* . . . : in these cases, participants always know immediately prior to the target word that the next word they encounter must be a target. If the artefactual account outlined here is correct, the *um* advantage should disappear. A number of claims have been made in the literature concerning the function of disfluency in conversational speech, both from the perspective of the speaker (who according to some accounts is conveying a message via disfluency) and from that of the listener, who may be able to make use of disfluency to glean information about the speaker's current status. Unfortunately, many of these claims are poorly supported by the existing evidence.

Mixdorff

Speaking styles – spontaneous and read = intonation and timing

Text-based and Signal-based Prediction of Break Indices and Pause Durations

Hartmut R. Pfitzinger & Uwe D. Reichel

p/1

Mean values and standard deviations of pauses at sentence boundaries were 716 ms and 336 ms, respectively, and of pauses within sentences were 327 ms and 132 ms, respectively